RESEARCH REPORT SERIES
(Survey Methodology #2023-03)

Use of Mobile Phone Location Data in Official Statistics, Social, Demographic and Health Studies

Elizabeth Nichols
Angela O'Brien
Shelley Feuer
Jennifer Childs

Center for Behavioral Science Methods
Research and Methodology Directorate
U.S. Census Bureau
Washington, D.C. 20233

Report issued: March 7, 2023

Disclaimer: This report is released to inform interested parties of research and to encourage discussion. The views expressed are those of the authors and not those of the U.S. Census Bureau.

Abstract

With high ownership and widespread daily use of mobile phones across the globe, governments are investigating whether mobile phone location data can be used to produce official statistics and generate information for the public good. Mobile phone location data appeals to many researchers due to its substantial coverage, low respondent burden, and potential for high accuracy. However, there are concerns about the legal and ethical implications of using these data, including the maintenance of privacy and provisions against government misuse. This paper provides an overview of different types of mobile phone location data, their uses, successes, and missteps by both government and private entities across the world.

We conducted a literature review to assess the state of international research in using mobile phone location data for official statistics and civic planning. We first searched for peer-reviewed articles detailing how governments in the U.S. and abroad use mobile phone location data for official statistics. We then expanded the search to include pilot studies, including studies by both United States government agencies, international governmental bodies, and other governments and groups outside the U.S. Finally, we extended the subject matter to include how both national and international governments use mobile phone data to trace and contain the COVID-19 pandemic. We begin with a description of the types of location data that mobile phones can provide and a brief discussion of the tension between these data and privacy concerns. We end with a summary of the literature.

Keywords: mobile phone data, official statistics

Suggested Citation: Elizabeth Nichols, Angela O'Brien, Shelley Feuer, Jennifer Childs. **(2023). Use of Mobile Phone Location Data in Official Statistics, Social, Demographic and Health Studies.** *Research and Methodology Directorate, Center for Behavioral Science Methods Research Report Series (Survey Methodology #2023-03).* U.S. Census Bureau. Available online at <https://www.census.gov/library/working-papers/2023/adrm/rsm2023-03.html>

Introduction

With high ownership and widespread daily use of mobile phones across the globe, governments are investigating whether mobile phone location data can be used to produce official statistics and generate information for the public good. Mobile phone location data appeals to many researchers due to its substantial coverage, low respondent burden, and potential for high accuracy. However, there are concerns about the legal and ethical implications of using these data, including the maintenance of privacy and provisions against government misuse. This paper provides an overview of different types of mobile phone location data, their uses, successes, and missteps by both government and private entities across the world.

We conducted a literature review to assess the state of international research in using mobile phone location data for official statistics and civic planning. We first searched for peer-reviewed articles detailing how governments in the U.S. and abroad use mobile phone location data for official statistics. We then expanded the search to include pilot studies, including studies by both United States government agencies, international governmental bodies, and other governments and groups outside the U.S. Finally, we extended the subject matter to include how both national and international governments use mobile phone data to trace and contain the COVID-19 pandemic. We begin with a description of the types of location data that mobile phones can provide and a brief discussion of the tension between these data and privacy concerns. We end with a summary of the literature.

Review of the Literature

Using the University of Maryland electronic library, Google Scholar, and various online newspapers, we found articles detailing how governments use mobile phone location data in official statistics. These articles spanned topics ranging from pilot studies for official population counts, COVID-contract tracing, and the use of mobile phone data in official government travel statistics. The literature review was conducted in 2021 with a few additional papers added after that time.

Mobile location data description

The literature indicated two main types of location data: Call Detail Records (CDR) and Global Positioning System (GPS) data.

Call Detail Records (CDR) Data and other data originating from cell phone towers

Mobile phone companies deliver calls and texts by transmitting signals through a series of cell towers. Calls, texts, and connections to the internet made by mobile phones ping a local cell tower and are recorded by mobile phone companies. The records produced by the signals are called Call Detail Records (CDR) and are owned by mobile network operators (MNO). Users do not opt-in to provide these data—the MNOs collect these data automatically for billing purposes. The number of records associated with the phone is dependent on the user calling or texting to create the record.

MNOs also produce records by collecting passive data called cell site location information (CSLI) (Grassini and Dugheri, 2020). When cell towers signal or probe to monitor network performance, they create records when a phone is on, but not in use. In our review, we were unable to determine if all MNOs collect these passive data or what data are available in the carrier's record. But this process would create more records for each phone, since in theory the MNO would track the phone even when the user is not making a call or texting. As a result, the amount of passive data produced is enormous.

The CDR record includes source and destination phone numbers, time of the call, duration of the call, the base transceiver station (i.e., the telecommunications equipment on the cell tower, often called the base station), among other items. CDR does not include the content of the call or text.

The base station associated with the record depends on the geographic location of the phone and network traffic (Pastor-Escuredo, et al., 2014). Each base station has a geographic latitude and longitude. The geographic area the base station serves is called the cell. Thus, the location associated with a CDR is not a precise measure of the exact signal location for the phone, but rather where the base station is located. In areas where two or more cell tower service areas overlap, triangulation can be used if the cell phone uses different stations for different parts of the conversation. Estimates vary on the phone's exact location within the cell. Although location estimates in heavily populated city neighborhoods may be accurate within a few blocks due to triangulation, estimates in rural areas with fewer cell towers may be less accurate and vary by more than a mile with a maximum of 25 miles (Simmons, 2022).

Mobility data from CDR is simply movement between cell tower regions. Mobility can only be detected if the user travels outside the cell and a ping is made to a different base station. If a phone user travels within the same cell, the CDR cannot pick up that there was any change in location. Thus, CDR generally cannot track short trips (Streetlight, 2020). Calabrese et al. (2013) used a *superset* of CDR provided by a company called AirSage. AirSage's Wireless Signal Extraction technology uses triangulation with multiple cell towers to better pinpoint the location of the device. However, most articles we found used only the CDR data associated with the base station location. Technology is evolving to improve the accuracy of the CDR location data.

Typically, several MNOs serve a geographic area with each one having a share of the market. In most of the articles reviewed, researchers worked with one MNO and had available to them only the records associated with the subscribers of that MNO. Most of the research studies also used only the data for a particular time frame, that is, weeks or months, and frequently for a particular geographic region of the country, like a city. The number of records, even with these restrictions, is large.

Global Positioning System (GPS) Data

Another type of location data that can be collected from mobile phones is called Global Positioning System (GPS) data. GPS data are collected by some applications (for example, weather or mapping applications) and some websites. These apps and websites receive signals from satellites to determine the devices' location, velocity, and elevation. The apps collect users' locations when they are actively in operation (known as operating in the foreground). Depending on the settings, these apps may also collect user locations while operating in the background (i.e., when simply installed on a mobile phone but not actively in use). GPS data are more precise than CDR data. Predicted location can be within 3 meters if there is no blockage from mountains, tunnels, or large structures, and the data are owned by each individual application (Streetlight, 2020).

Users typically allow GPS location data to be collected by responding to notifications on their devices that ask the user whether they are willing to share their location data. If the user selects "allow," then the location data are collected. The way users are asked for their permission varies across apps and across brand of device. Some allow the user to share location services only once, always, or never by an app (see https://support.apple.com/en-ca/HT207092). Third-party brokers can buy these data and repackage and resell the information to be used by marketers or advertisers. There appears to be more research using CDR than GPS data, with researchers working directly with MNOs to acquire the data rather than third party providers.

Protecting privacy

While both types of data (CDR and GPS) are collected at the individual device level, nearly all the research projects we reviewed used anonymized and/or aggregated location data from mobile phones. Third-party providers do not provide raw, granular data. Anonymization removes all personal identifiers from the data, including the phone number, and aggregation typically refers to providing counts at various levels of geography and/or time periods rather than individual data records. These disclosure avoidance techniques are used to protect the privacy of where individuals live, work, or travel. While the anonymized techniques appear to be universally used, they do not appear to be foolproof as we see in the news (Valentino-DeVries, et al., 2018) where several "anonymized" phones were able to be linked to individuals based on where they lived and worked.

One article critiques the typical anonymization methods used with these data, stating that four data points containing time and general location is enough information to identify approximately 95% of phones in a dataset of 1.5 million people (de Montjoye et al., 2018). These authors call for an agreed upon framework for the use of anonymous mobile phone data. They then propose four framework models for *privacy-conscientious* use of mobile data, including limited release, pre-computed indicators and synthetic data, remote access, and question and answer. Limited release is a more traditional form of data sharing. The anonymized data are transformed through sampling and data coarsening (i.e., variable suppression), prior to release to another party under a legal contract. Pre-computed indicators are aggregated counts like the number of calls per cell phone tower over a certain time and synthetic data are generated by a model based on the actual mobile phone data. The authors claim that little work has been done to create or study synthetic mobile phone data. Remote access is where trusted parties perform the computation in a secure area associated with the data owner and the data owner has control over what data are released. With question and answer, the data stay with the data owner and the researcher can ask questions through a query system and receive the answer, but never manipulate the data themselves.

A report from the U.S. President's Council of Advisors on Science and Technology on Big Data (2014) acknowledges how challenges to privacy have emerged as new technologies make additional data available. The report suggests focusing on the policies around the uses of the data rather than methods for collection and analysis as technology changes so quickly that it is too easy to fall behind.

Uses of Location Data Derived from Mobile Phones for Official Statistics

A 2016 report from the Office of National Statistics (ONS) in the United Kingdom presents a literature review of international uses of mobile phone location data for official statistics. This report highlights the strengths of these data, including low respondent burden since the data are collected passively, along with the ubiquity of mobile phone coverage (93% of UK adults have a mobile phone). Additionally, because location and time data are collected in real time, location data are potentially more accurate than statistics coming from data collections that rely on respondents' memories. Finally, the report concludes that these data might generate insights into population behavior at much smaller geographies than traditional surveys because these data can be linked to areas ranging from 3 meters to 1 mile. ONS predicts that these data would best be suited for estimates that require data on mobility, such as for urban and transport planning and tourism statistics. They also suspect that mobile location data would be helpful for epidemiological research. Their literature review highlights the need for rules and oversight regarding personal data protection and data security, ethical concerns, commercial sensitivities and costs, and technical infrastructure. While some data uses such as measuring population density of an area are considered ethically sound, other uses such as inferring nationality or ethnicity based on the phone's language setting are more questionable. Their report acknowledges how working with the MNOs to acquire the data can place significant costs on the MNO and being able to accept the large datasets also puts strain on the technical infrastructure of the receiving organization. Furthermore, the ONS report urges future studies on potential bias associated with mobile phone location data, such as overestimating population density since some individuals have more than one phone.

In 2019, the United Nations (UN) Global Working Group on Big Data published a report detailing how countries are testing and using mobile phone location data for official statistics (United Nations, 2019), many of which are detailed in the current report. The UN encourages additional research using mobile phone location data as the technology could be particularly helpful in countries lacking the resources for more labor-intensive data collections such as surveys. Prior to that report, a UN report focusing on the Caribbean encourages the use of big data for official statistics and cites tourism statistics from Estonia as one example of mobile data use (Abdulkadri, et al., 2016).

The remaining sections of this report summarize the articles we found documenting specific studies using mobile phone data. We provide an overview of the articles we found by topic. Included are the goal of the research, the country where it was conducted, the year, the type of data (if it was clearly defined), and the results.

Currently only two countries, Estonia and Indonesia, use these data for official statistics. The remaining articles were pilot and research studies conducted by government entities, academics, and private organizations. These groups pilot tested how mobile phone location data could be used for a variety of purposes such as studying population density, mobility, and tourism. There were also a few studies which investigated whether these data could be used to predict socioeconomic characteristics, future crime areas, or establishment sales. The COVID pandemic in 2020 appeared to accelerate this type of research. Numerous articles were devoted to attempting to track the virus to predict where resources would be needed.

Production of official statistics with mobile phone location data

Since 2008, the central bank of Estonia has worked with Estonian MNOs to use mobile phone location data for travel statistics, allowing researchers to see how often and for what duration of time people spent traveling to and from Estonia. Estonia also uses that same data along with credit card data to quantify trade in service (United Nations, 2019).

Indonesia uses mobile phone location data to measure inbound cross-border visitor arrival statistics in areas that do not have immigration checks. Grassini and Dugheri (2020) document these and other efforts to measure tourism statistics and conclude that while the data hold promise compared to traditional measures (such as accommodation data), costs and standard definitions of what defines a tourist using these data are still of concern.

Research studies and COVID-19 tracing

While the list of official statistics that are derived from mobile phone location data is small, there are many research studies that have used these data. Academics, government agencies, and other groups have conducted a variety of pilot studies using mobile phone location data. The topics of these pilot studies center on some common themes. Some articles describe how these data were used to study population estimates of either a geographic region or of inbound migration (people from one country in another country). Mobility (or lack of mobility) within a country or geographic region, such as for commuting purposes, natural disasters, epidemics, or because of government regulations, is another study topic, with numerous articles around COVID-19 tracing. The third topic that appears with some frequency is determining whether mobility is related to socio-economic factors, such as poverty. Finally, there are a few pilot studies that examined whether mobility data could be used to improve existing models to estimate such things as exposure to air pollution or prediction of crime areas.

The following sections detail these pilot studies. When reading through the examples it becomes clear that many of the articles highlight studies conducted outside of the U.S. Within the U.S. there has been some research about using these data for regional traffic studies or commuting patterns (Colak et al., 2015), but the number of studies is much smaller than for other counties. We suspect that there is more interest determining whether mobile phone location data can provide information to make timely and cost-effective decisions in emerging economies that do not have the resources for ongoing survey data collections.

Population estimates

In our review, many researchers are interested in whether mobile location data can be used to create population estimates of an area. Cell phone coverage is an issue of concern when making population estimates. Low cell phone penetration rates underestimate the population, but multiple cell phones per person overestimate the population as would multiple records associated with the same phone (Deville et al., 2014). Calibration against a benchmark is needed to readjust the estimates from these data (Douglas et al., 2015). Several examples of population count research follow.

- A partnership between Statistics Belgium and Eurostat resulted in a comparison of CDR to actual population data from the 2011 Belgium Census. The researchers demonstrated consistency between the two datasets (De Meersman et al., 2016).
- Similarly, Deville and colleagues (2014) compare national census population data for Portugal and France to population estimates using CDR and conclude the CDR results successfully match the other sources.
- Researchers from University of California examined whether they could measure the population in Milan, Italy with a regression model examining the relationship between call volume (using CDR data) and census population numbers. They found positive results with the model, meaning it was feasible, but with some caveats. The model would need to be adjusted for each local community and recalibrated periodically with population counts in key regions from a truth source, like a census (Douglas et al., 2015).
- Researchers from three separate universities used CDR from one MNO in 2011 and 2012 to measure population in a region of the Ivory Coast. They compared their estimates to census and satellite imagery estimates from The AfriPop Project. Their analyses resulted in a larger population estimate in urban areas and lower population estimates in the rural areas with the

CDR compared to AfriPop data. They attribute some of the differences to differential mobile phone penetration rates in rural areas and recent economic development in one more urban region which the CDR identified as having higher population than the AfriPop source. The authors conclude that the mobile phone location data helps identify population distributions in urban areas (Sterly et al., 2013).

- Researchers used mobile device data (navigational-GPS data) and analysis tools from StreetLight Data, Inc. to estimate visitation in parks and protected areas in California, in the United States. The results were generally positive, but the authors note that calibration techniques are needed, and the results should be considered estimates and not direct counts (Monz et al. 2019).
- The Environmental Protection Agency (EPA), a U.S. federal agency, used a commercially available dataset from AirSage, Inc. that contained counts from both CDR and GPS data to study visits to water recreational areas in a pilot study (Merrill et al., 2020). AirSage, Inc. is a third-party provider of anonymized and aggregated data. The EPA had over 500 areas of interest in the New England area and used data from the summer of 2017 (June through September). They compared these data to standard observational visit counts collected from those areas (such as from entrance fees, onsite counts of small access points, and visitation estimates for managed beaches). The mobile phone data resulted in overestimating the observation counts. They too conclude that calibration and validation against another source, such as observational reports, is needed to use these data with confidence.
- Researchers at the University of South Carolina and the University of Arkansas examined changes in visitation patterns at six U.S. national parks in response to the COVID-19 pandemic. They used data collected anonymously from mobile devices, provided as part of SafeGraph's Social Distancing Metric dataset. They compared these data to official visit counts from the U.S. National Park Service. They conclude that mobility records can effectively capture park visitation patterns and origin of visit, which can help with park planning needs (Kupfer et al., 2021).
- In 2013, Statistics Netherlands conducted a study of inbound tourism using CDR from one MNO. They found the CDR estimates were much larger than tourism accommodation statistics. They attribute the difference to the fact that accommodation statistics do not count any foreign visitors who stay with friends or family and accommodations with less than 10 beds do not have to specify how many of their guests were from out of the country. The results from the mobile phone location data were matched the geography where the record was made. The resulting locations matched general behavior of tourists on specific holidays, special events like the Union of European Football Associations (UEFA) cup football game, and in tourism spots, like the coast, where peaks in population occurred. The researchers conclude that CDR has broad and feasible potential for tourism statistics (Heerschap et al., 2014).
- The Statistical Office of the Metropolitan City of Florence, Italy conducted a study using CDR to study the strengths and weakness of the counts of visitors based on the phone data compared to official accommodation data. They used data from one MNO over a five-month period in 2016. While they found similar results for foreign visitors, measures of Italian visitors to Florence were higher in CDR than in the official accommodation data. They attribute the differences to difficulty defining who a visitor is with the CDR data, such as including commuter workers who spend the night in Florence less than three times a week or individuals who went on holiday

during the summer months (Grassini and Dugheri, 2020). This article also included an extensive review of the literature on use of mobile location data for tourism statistics.

Mobility

Mobile phone location data is also used to study geographic mobility. To study geographic mobility, researchers need to identify a user's "origin" (often referred to as the "home" location) and destination. However, MNOs do not provide either to researchers, leading Vanhoof and colleagues (2018) to call for more research on detecting home locations. To support this effort, Vanhoof and colleagues conducted analyses on CDR data from around 18 million users in France. They compared the accuracy of five different algorithms for detecting a user's home location. According to these researchers, the algorithms frequently made flawed home predictions. The accuracy of home detection is a limitation in all the following studies where the researchers selected a home location algorithm to use and examine mobility as measured by mobile phone location data.

- In one article, researchers explored the use of CDRs to generate origin – destination (O-D) trips of various purposes (home, work, other) and different times of day. They assigned O-D information for a sample of users in two cities, Boston and Rio de Janeiro, using large CDR data sets. They compared the results to validation data from the census and travel surveys for Boston and concluded that the results using CDR data could accurately estimate these intertown trips in Boston and inter-district trips in Rio de Janeiro (Colak et al., 2015).
- Zhou et al. (2021) investigated whether mobile phone location data can help with urban planning to minimize commute times between residences and employment. They used CDR for two weeks in October 2014, from one MNO for Shanghai, China. The authors used algorithms from Calabrese et al (2013) and Zhou et al. (2018) to identify home and work locations from these data. The authors found that commute times are reduced with compact urban design, where commercial and residential area were mixed. They found industrial areas were separate areas in Shanghai and increased commute time to those areas. They recommend that industrial areas be mixed in with commercial and residential areas (Zhou et al., 2021).
- Another case-study in China examined whether land consolidation rules in China affected the lifestyle of residents as measured by their movement in those areas. CDR data were obtained from one MNO for one city in a rural area of China in 2017. The researchers measured mobility of individuals in the rural parts of that area and other more nonrural areas. They concluded that land consolidation rules did not necessarily harm the lifestyle of rural residents. While the commute time from home to farm increased the commute time for other activities decreased with an average decrease in time spent traveling (Liu et al., 2020).
- Researchers used CDR to measure mobility in and around Morristown, NJ for two months to determine the feasibility of use among urban planners. They compared the data to the U.S. Census data and found consistency. The research showed that they could group city dwellers and visitors into categories based on mobile phone usage patterns. When they combined the clustering analysis results with contour maps, researchers showed how people who lived outside of the city were not as likely to come back into the city over the weekend. These types of insights are of value to urban planners (Becker et al., 2011).
- Researchers used CDR data to study mobility (both vehicular and nonvehicular) in the Boston, Massachusetts area. These researchers compared their estimates of home location to the 2020

Census population counts for eight selected counties. They also compared the data to annual vehicle safety inspection records including odometer readings to create annual travel distance. The researchers conclude that the timely CDR data could potentially be used to complement more traditional travel surveys (Calabrese et al., 2013).

- Between 2010 and 2014, researchers used CDR to study how internal migration patterns in Namibia change over time. They used data from one MNO with 76% of the market share. They conclude that their use of CDR data complemented the national migration statistics and appeared very promising for timely and local data (Lai et al., 2019).
- Researchers investigated the migration patterns before and after the 2010 earthquake in Haiti using CDR from 1.9 million people. They were able to map time of mobility, distance traveled, and duration stayed at the destination location. They concluded that approximately 23% of the population of Port-au-Prince left the city after the earthquake and they also concluded that those who left within three weeks after the disaster traveled to places that they had traveled to before the earthquake (Lu et al., 2012).
- In 2011 and 2012, Statistics Netherlands conducted two sequential studies using GPS data collected through a downloadable app for mobility statistics in the Netherlands. The app registered location and time data for the sample of respondents. Researchers compared these data to survey data and found the survey data overestimated travel. There were limitations with downloading an app since not everyone wanted to participate and some of them faced technical challenges with the download. The article did not discuss the percent who declined participation for either of those two reasons in their two studies (Heerschap et al., 2014).
- The Office of National Statistics (ONS) directed a study of commuter flow data using CDR from one MNO (Vodafone) over four weeks throughout March and April 2016 in three London boroughs. They then compared the results to the census data collected in 2011. While there was a correlation between the two data sources, it was difficult to distinguish home workers or commuters who travel very short distances. Therefore, it's possible that other groups of people, such as students or those visiting a nearby shopping area twice a week, were mistaken for workers (BBC News, 2017).
- The ONS travel and tourism team compared CDR data to survey results from the International Passenger Survey. They evaluated estimates of both foreign residents visiting the United Kingdom (UK) and UK residents traveling abroad. Estimates of inbound visitors were greater for the survey data than for the mobile data, while estimates were closer for outbound travelers. ONS plans to continue this type of research with mobile data (Data Science Campus, 2021).
- The National Academy of Sciences, Engineering, and Medicine (2021) published a review of the use of GPS and other data for travel patterns and traffic analysis in the U.S. Nearly every state in the U.S. responded to their request for information. States use mobile phone data and other data sources to study causes of traffic and effects on traffic due to incidents, such as a bridge being closed.

Tracking other social and economic patterns

We found a few studies that investigated whether there is a relationship between CDR and population behavior. We also found some studies that looked at whether location data from mobile phones could be used to improve existing models for pollution exposure, crime locations, and other socio-economic indicators.

- Using CDR from one MNO, researchers studied segregation of ethnicity groups in Estonia's capital city. In addition to the CDR, the MNO provided the researchers with the gender, year of birth, and user language selected to communicate with the MNO. A sample of over 5,000 individuals was selected from the CDR data with 54% percent of the sample being Estonian-speaking Estonians and 46% Russian-speaking Estonians. Using language as a proxy for ethnicity, they studied the mobility of the sample and conclude that ethnic groups in Estonia interact with each other more during the day than the evening, and their interaction rates differ by season (Silm et al., 2014).
- Pastor-Escuredo and colleagues (2014) investigated how the volume of CDR data is affected by natural disasters. They conducted a retrospective study using CDR data to examine call volume in the region of Tabasco, Mexico around the time of the 2009 floods. Results indicated that mobile phone activity (simply the number of calls or texts) in that region was unusually high when the emergency was declared as opposed to when the warning was made. Future use of these data for emergency response was proposed, with a special focus on areas that do not heed warnings.
- In a study conducted in Rwanda using survey data from 2009 and CDR data from 2010, Blumenstock and colleagues (2015) created models using mobile phone usage patterns to predict socioeconomic status. They concluded that these models can accurately predict wealth distributions at local levels in a cost-effective way. They also suggest that the method is especially useful in developing countries that do not have the resources for ongoing surveys or censuses.
- Another study also investigated the relationship between CDR data and poverty. This research was conducted by the World Bank and focused on whether CDR data could be used to predict poverty estimates in Guatemala. Hernandez and colleagues (2017) used benchmark poverty estimates from surveys and censuses. They determined that CDR poverty estimates from models were more accurate in urban areas than in rural areas, which they attributed to the greater number of cell towers and the larger population in urban areas.
- Song et al. (2019) investigated whether CDR data could be used to predict where crime occurs and whether these mobile phone location data were better at predicting crime areas than other models which used likely places where crime occurs where large numbers of people congregate. Using CDR data from an MNO with 23% of the market share in one city in China, the researchers found that the probability of crime occurring increases in areas where the offender lives and where large numbers of people are located. This relationship between populations size and proximity to the offender was stronger than simply identifying the likely places where crime occur, such as subway stations (Song et al., 2019; Paul & Brantingham, 1995).
- In another study, researchers used a CDR dataset containing 4.6 million records for approximately 10,000 phones in Shenzhen, China to study air pollution exposure. Results show that the CDR based approach (tracking where people moved and creating pollution exposure estimates from those locations) can differ substantially from home location-based pollution metrics (where someone's exposure estimate is only based on one location). They conclude that the CDR based method could be used to improve future air pollution epidemiology studies (Yu et al., 2018).
- Researchers used a CDR dataset from 17 months in 2009 and 2010 in an "emerging economy" Latin American country. They aimed to determine whether calling variables (e.g., number of

calls and call length) and the mobility of individuals as calculated by CDR data can be used to predict future socioeconomic indicators such as total assets, total number of employed citizens, etc. They compared these data to the socioeconomic time series computed by the local National Statistical Institute. The results show that using a multivariate time-series computed with different sets of usage and mobility variables yield reliable predictive results (Frias-Martinez et al., 2013).

- In South Korea, researchers conducted a study with five participants who downloaded an app that would track GPS signals. Examining the locations visited, the researchers were able to determine occupancy status of residential units using those data (Kim, et al. 2021).

Public health and epidemiology tracing

The most recent widespread use of mobile phone location data has come in response to the COVID-19 pandemic (Benjamins, Vos, Verhulst, 2021). There was a great interest in whether mobility was reduced due to the stay-at-home orders or whether the public relocated from urban areas during the pandemic. This information allowed governments to predict where resources, including health and economic resources, needed to be directed.

Other governments used location data to provide warnings to the public about possible exposure. The type of data acquired to monitor mobility and provide warnings about exposure also varied across countries. Several COVID-tracing apps used Bluetooth technology, which is distinct from CDR or GPS location data. These apps track the Bluetooth keys of phones that have been in close proximity. If a user receives a positive COVID-test and uploads that information into the app, then the app shares an exposure notification with all the Bluetooth keys the app has stored while keeping who was positive concealed. While it might seem like the app tracks user locations; this technology does not. However, the success of the technology depends on the user initiating actions such as downloading and installing the app on a phone and then, if the user ever tested positive, to enter those results in the app.

The Australian government created a COVIDSafe application using this Bluetooth technology (Cha, 2020). The Interfederal Committee for Testing and Tracing in Belgium created a voluntary contact tracing app using this technology called Coronalert. This app was downloaded by less than 30% of the Belgium population and the opinion of some medical personnel is that it did not help prevent the spread of COVID (Chini, 2021). Singapore developed the Bluetooth technology to develop a contact tracing app called “TraceTogether.” The article stated that the app was not widely used in Singapore. Although users were encouraged to upload positive COVID-19 test results to the app, users had to consent to share their data with the government before uploading. The app also automatically erased data from users’ phones after 14-16 days (Cha, 2020). Twenty-five U.S. states and the District of Columbia promoted the voluntary use of COVID-tracing apps that implement the Bluetooth technology. State populations covered by these apps range between 1% to 46%. It is unclear whether these apps reduced or slowed the spread of COVID in the U.S. (Ladyzhets, 2021).

While the Bluetooth technology did not track locations visited, Cha (2020) discusses how some countries, mostly in Asia, appear to combat the pandemic by adopting high tech public tracing methods that emphasize the public good over individual privacy. "Asian societies shifted in the direction of treating these technologies as contributing to a public good that is well worth the temporary and necessary incursions of privacy" (p. 12).

- In Hong Kong, Cha reports that the government worked with private software companies to construct digital, real-time apps to alert citizens to movements of disembarked passengers from a cruise ship with high COVID-19 rates.
- In February 2020, the government of South Korea allowed tracking individuals' phones to create a publicly available map so that citizens could check whether they may have crossed paths with any coronavirus patients. South Korean government proactively sent regional text messages warning people they may have come into contact with someone carrying the virus, with specific details about the individual's travel location. The director of the Korea Centers for Disease Control and Prevention recognized that the site infringes on civil liberties, "It is true that public interests tend to be emphasized more than human rights of individuals when dealing with diseases that can infect others." (Cha, 2020).
- In Taiwan, the government tracked mobile phone data and alerts authorities when someone who is supposed to be quarantined leaves using an "electric fence" system. Even turned off phones could result in a police visit (Hamilton, 2020; Cha, 2020).

However, other countries also used these types of apps for monitoring purposes.

- In March 2020, the Iranian government authorized a coronavirus diagnosis app that collected users' real-time location data. The government sent a message to millions of Iranian citizens telling them to install the app, AC19, before going to a doctor's office or a hospital. The app was said to be able to diagnose the user with coronavirus by asking a series of yes or no questions. Due to this misleading claim, the app was removed from the Google Play store in March 2020 (Hamilton, 2020).
- Between March 17 and April 22, 2020, Israel's Security Agency no longer had to obtain a court order to track individuals' phones as part of a set of COVID measures approved by Prime Minister Benjamin Netanyahu. The law also required that all data collected must be deleted after 30 days. They collected data to identify people who have crossed paths with known patients. A text was sent then to individuals to warn them that they might have been in contact with a COVID-positive individual (Hamilton, 2020). There were ongoing court challenges to these measures by rights groups and the Israeli parliament suspended the use of these data after a month (Williams & Rabinovitch, 2020).
- The Polish government used a home quarantine app to ensure people who are supposed to be quarantining themselves for 14 days stay at home. The app also requests periodically for the person to send a geo-located selfie to the government with the understanding that the police will be alerted if the user fails to comply within 20 minutes (Hamilton, 2020).

Other countries use aggregate and anonymous data to measure mobility to help direct resources, determine whether mobility declined, or estimate how populations shift across the country. Google has created anonymous and aggregated mobility reports by country, and within the U.S., by state and region with downloadable data (Google, 2021).

- One major mobile phone network shared anonymized location data with the Austrian government. That government then used that anonymized data to map population movements (Hamilton, 2020).
- Between March and May 2020, the French National Institute of Statistics and Economic Studies (INSEE) collaborated with three of the four MNOs operating in France to produce statistics using

aggregated and anonymous counting indicators from network signal data to study population movements and mobility due to COVID-19 and the government's responses (Coudin et al., 2021).

- The Public Utilities Regulatory Authority (PURA) and The Gambia Bureau of Statistics used anonymized and aggregated CDR to show how the Gabian population moved from the capital city region to rural areas during March through May 2020 during the COVID-19 lockdown (Arai, et al., 2021). Two out of four MNOs provided anonymized data to PURA.
- Between March and April 2021, the Ministry of Health and Emergency Operations Centre in Malawi used anonymized and aggregate CDR data to show mobility of individuals. Online dashboards were created to quantify and locate mass gatherings and travel between subdistricts. These dashboards helped identify possible chains of spread, irregular high-volume movements, and physical interconnectedness between regions (Green, et al., 2021).
- Nigeria Governors' Forum instigated a collaboration with Nigeria's largest MNO to predict worse-case COVID-19 infection rates and determine where to focus economic relief. They would now like to have an ongoing relationship with the MNO to use data-driven decisions for the public good (Gilbert, et al., 2021).
- Public Health Agency in Sweden successfully collaborated with a Nordic Telecoms operator to use aggregate and anonymized data to study the mobility of the public. The results indicated that travel decreased during the pandemic as recommended by the Swedish government. The article highlights the collaboration between a telecom operator and a public institution, the use of mobility data in evidence-based policymaking, and the adherence to personal data protection laws (Ågren, et al., 2021).
- Statistics Sweden collaborated with a large telecom operator, Telia, in 2020 to study the feasibility of using cell phone data to study population mobility and public/private partnerships using these data (Vlag, et al., 2022). Results indicate that the population shifts in the summer from cities and suburbs to the countryside.

However, during this time there was also more scrutiny over acquiring and using these data. For example, the Center for Disease Control and Prevention (CDC) in the U.S. acquired anonymized and aggregated mobile phone data from SafeGraph. Twenty-one use cases for these data were detailed in the plan; most of them were associated with COVID-19 pandemic monitoring. The article highlights that some of the use cases were not related to the pandemic but rather geared toward studying other public health issues. The article cites general concern about the scope of the request. While it does not document any disclosure of individual identities from the data source, it does indirectly suggest privacy concerns by mentioning that the data supplier has been banned from the Google play store (Cox 2022).

U.S. government research of mobile phone location data

The U.S. federal government has yet to fully explore the potential utility of mobile phone location for official statistics, but as Jarmin (2019) documents, government statisticians are investigating a range of new data sources to improve statistics. As one of the few examples of using mobile phone data, staff at the Environmental Protection Agency published a study detailing their pilot test of using aggregated and anonymized mobile location data in predicting visits to water-recreational areas in New England (Merrill et al., 2020). After calibrating and validating the data against actual visits counts, they found promise in the models using these location data for estimating daily visits.

The U.S. Census Bureau currently does not use mobile phone location data for estimation or evaluation purposes. However, there is interest in determining whether these data could be used to supplement travel studies or whether there are relationships between these data and other estimates. Originally, some Census Bureau staff thought that CDR data might be able to help identify places of residence or the occupancy status of these residences. Having an indicator of occupancy would be particularly helpful in more rural areas where sending enumerators to check occupancy status is costly. However, the granularity needed to answer those questions is not available in the typical aggregated data sources. More importantly, the location associated with CDR data is not accurate enough for a specific latitude and longitude. The CDR data could possibly be used to compliment the LEHD data that shows origin-destination for employment or shopping areas. It also could be useful if measuring mobility due to disasters or transportation. These data would be more useful if they could be linked to demographic data, but that would have to be done in a confidential computing environment where the individual data could not be observed by analysts creating the statistics.

To help predict occupancy or residential units, GPS data would be needed as that data is far more accurate. Additionally, we found two WiFi signal research projects which appear to be able to generate results at a very small geographic level. These might be useful avenues to pursue for any in-person field operation measuring occupancy.

- Instead of using CDR or GPS, an academic research project used Wi-Fi probe requests generated by mobile devices in a geographic area to predict retail food establishment sales in the UK using a statistical model. These data were collected by a large network of sensors installed across 105 towns and cities in the UK. The models using the WiFi probe data more accurately predicted sales than models not using the data, and the improvement differed between fast-food restaurants locations and family restaurant locations (Trasberg et al., 2021).
- In Japan, researchers used ambient (i.e., ever-present) WiFi signals from homes to reduce the number of residential addresses that had to be visited to determine occupancy status (Konomi et al., 2019).

Summary

This report documents the widespread use of mobile phone location data in pilot studies across the globe to measure geographic mobility, travel, population density, residential structures, and occupancy status. In a few studies, researchers explored whether these data can be a proxy to measure other population characteristics such as poverty. Although many studies report successful outcomes with calibration adjustments, others suggest more rigorous analysis should be conducted to confirm that the findings are unbiased. Use of these data grew considerably during the COVID-19 pandemic, often to measure compliance with mobility restrictions and for resource planning purposes.

Acquiring the data to perform the analysis is the first step. Several papers document the emerging partnerships between governments (especially European and Latin American governments) and MNOs and the challenges of data privacy, security, and commercial sensitivity that emerge with these relationships (De Alarcon et al., 2021; Vespe, et al., 2021). Milusheva et al. (2021) documents the attempts of 41 developing countries to build relationships between the owners of the mobile data and governments that protects privacy and maintains financing for the use of these data. Nine of the 41 were successful. Jansen and colleagues (2021) document how Estonia, Ghana, and Gambia have used mobile network data related to five principles associated with mobile location data use that they identify. Those principles are necessity and proportionality, professional independence, privacy protection, commitment to quality, and international comparability.

The articles make clear that the relationship between the researcher and the data provider is critical to being able to obtain the data. Guidelines indicating how the data will be used and protected are necessary. Many of the articles emphasized the importance of using anonymized and aggregated data to protect individual privacy, but concern that even with those procedures in place the data might identify individuals as technology advances. Studies that measure the public's attitudes towards governments using these data appear to be limited, particularly in the U.S. The recent newsbreak expressing concern that the CDC has acquired anonymized and aggregate mobile phone data for research purposes confirms that at least in some groups, there is wariness towards use of these data by the U.S. government (Cox 2022). Additional research is needed to measure the public's attitude toward the U.S. federal government using these data, what are acceptable use cases and what are not.

References

Abdulkadri, A., Evans, A., & Ash, T. (2016). An assessment of big data for official statistics in the Caribbean: Challenges and opportunities. *ECLAC Subregional Headquarters for the Caribbean*.

Ågren, K., Bjelkmar, P., & Allison, E. (2021). The use of anonymized and aggregated telecom mobility data by a public health agency during the COVID-19 pandemic: Learnings from both the operator and agency perspective. *Data & Policy*, *3*(E17). doi:10.1017/dap.2021.11

Arai, A., Knippenberg, E., Meyer, M., & Witayangkurn, A. (2021). The hidden potential of call detail records in The Gambia. *Data & Policy*, *3*(E9). doi:10.1017/dap.2021.7

BBC News. (2017, November 7). Census "Could use mobile phone data instead of questions." *BBC News.* www.bbc.com/news/uk-41898318

Benjamins, R., Vos, J., & Verhulst, S. (2021, Oct 6). *Mobile Big Data in the fight against COVID-19.* Data^ Policy Blog. https://medium.com/data-policy/mobile-big-data-in-the-fight-against-covid-19-6e550d4f2bf8

Becker, R. A., Caceres, R., Hanson, K., Loh, J. M., Urbanek, S., Varshavsky, A., & Volinsky, C. (2011). A tale of one city: Using cellular network data for urban planning. *IEEE Pervasive Computing*, *10*(4), 18-26.

Blumenstock, J., Cadamuro, G., On, R. (2015). Predicting poverty and wealth from mobile phone metadata. *Science* 350, 1073–1076.

Brantingham, P. & Brantingham, P. (1995). Criminality of place: crime generators and crime attractors. *European Journal of Criminal Policy and Research*, 3(3), 5-26.

Calabrese, F., Diao, M., Di Lorenzo, G., Ferreira Jr, J., & Ratti, C. (2013). Understanding individual mobility patterns from urban sensing data: A mobile phone trace example*. Transportation research part C: emerging technologies, 26*, 301-313.

Cha, V. (2020). Asia's COVID-19 lessons for the West: Public goods, privacy, and social tagging. *The Washington Quarterly*, 1-18.

Chini, Maïthé. (2021). Over 11 million Covid Certificates for travel downloaded in Belgium. The Brussels Times. https://www.brusselstimes.com/news/belgium-all-news/179917/over-11-million-covid-certificates-for-travel-downloaded-in-belgium.

Colak S, Alexander LP, Alvim BG, Mehndiratta SR, Gonza´lez MC. (2015). Analyzing cell phone location data for urban travel: current methods, limitations, and opportunities. Transportation Research Record. 2526(1):126–35.

Coudin, E., Poulhes, M., & Suarez Castillo, M. (2021). The French official statistics strategy: Combining signaling data from various mobile network operators for documenting COVID-19 crisis effects on population movements and economic outlook. *Data & Policy*, *3*(E10). doi:10.1017/dap.2021.1

Cox, J. (2022, May 3) *CDC Tracked Millions of Phones to See If Americans Followed COVID Lockdown Orders*. Vice. https://www.vice.com/en/article/m7vymn/cdc-tracked-phones-location-data-curfews

Data Science Campus. (2021, October 5). Using Mobile Phone Data for Enhancing International Passenger survey Traveller Statistics. https://datasciencecampus.ons.gov.uk/using-mobile-phone-data-for-enhancing-ips-traveller-statistics/

De Alarcon, P., Salevsky, A., Gheti-Kao, D., Rosalen, W., Duarte, M., Cuervo, C., Muñoz, J., Pascual, J.M., Schurig, M., Treß, T., de la Cuesta, C., & Frias-Martinez, E. (2021). The contribution of telco data to fight the COVID-19 pandemic: Experience of Telefonica throughout its footprint. *Data & Policy, 3*(E7). doi:10.1017/dap.2021.6

De Meersman, F., Seynaeve, G., Debusschere, M., Lusyne, P., Dewitte, P., Baeyens, Y., Wirthmann, A., Demunter, C., Reis, F., & Reuters, H. (2016). Assessing the quality of mobile phone data as a source of statistics. *European Conference on quality in official statistics* (Q2016).

de Montjoye, Y.A., Gambs, S., Blondel, V., et al. (2018). On the privacy-conscientious use of mobile phone data. Sci. Data. 5:180286 doi: 10.1038/sdata.2018.286

Deville, P., Linard, C., Martin, S., Gilbert, M., Stevens, F.R., Gaughan, A.E., Blondel, V.D. & Tatem, A.J. (2014). "Dynamic Population Mapping Using Mobile Phone Data." Proceedings of the National Academy of Sciences 111(45): 15888–15893. Doi:http://dx.doi.org/10.1073/pnas.1408439111.

Frias-Martinez V, Soguero-Ruiz C, Frias-Martinez E, Josephidou M. (2013). Forecasting Socioeconomic Trends With Cell Phone Records. In Proceedings of the 3rd ACM Symposium on Computing for Development, 1–10.

Gilbert, J., Adekanmbi, O., & Harrison, C. (2021). Using mobile big data to support emergency preparedness and address economically vulnerable communities during the COVID-19 pandemic in Nigeria. *Data & Policy, 3*(E21). doi:10.1017/dap.2021.12

Grassini, L., & Dugheri, G. (2021). Mobile phone data and tourism statistics: a broken promise?. *National Accounting Review*, *3*(1), 50-68.

Green, D., Moszczynski, M., Asbah, S., Morgan, C., Klyn, B., Foutry, G., S., Ndira, N. Selman, M., Monawe, A., Likaka, R., Sibande, & T., Smith. (2021). Using mobile phone data for epidemic response in low resource settings—A case study of COVID-19 in Malawi. *Data & Policy, 3*(E19). doi:10.1017/dap.2021.14

Google. COVID-19 Community Mobility Reports. Last accessed 12/1/2021 https://www.google.com/covid19/mobility/.

Hamilton, I. (2020, April 14). Compulsory selfies and contact-tracing: Authorities everywhere are using smartphones to track the coronavirus, and it's part of a massive increase in global surveillance. *Business Insider.* www.businessinsider.com/countries-tracking-citizens-phones-coronavirus-2020-3

Heerschap, N., Ortega, S., Priem, A., & Offermans, M. (2014). Innovation of tourism statistics through the use of new big data sources. In 12th global forum on tourism statistics, Prague, CZ (Vol. 716).

Hernandez, M., Hong, L., Frias-Martinez, V., Whitby, A., & Frias-Martinez, E. (2017). Estimating poverty using cell phone data: Evidence from Guatemala. *World Bank Policy Research Working Paper*, 7969.

Jansen, R., Kovacs, K., Esko, S., Saluveer, E., Sõstra, K., Bengtsson, L., Li, T., Adewole, W.A., Nester, J., Arai, A., & Magpantay, E. (2021). Guiding principles to maintain public trust in the use of mobile operator data for policy purposes. *Data & Policy*, *3*(E24). doi:10.1017/dap.2021.21

Jarmin, R. S. (2019). Evolving Measurement for an Evolving Economy: Thoughts on 21st Century US Economic Statistics. *Journal of Economic Perspectives*, 33 (1): 165-84.

Kim, S., Lee, E., Park S. & Kwon, K. (2021). Non-intrusive Occupancy Estimation using a Smartphone GPS, *International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 1853-1855, doi: 10.1109/ICTC52510.2021.9621040.

Konomi, S. I., Sasao, T., Hosio, S., & Sezaki, K. (2019). Using ambient WiFi signals to find occupied and vacant houses in local communities. *Journal of Ambient Intelligence and Humanized Computing, 10*(2), 779-789.

Kupfer, John A., Zhenlong Li, Huan Ning, and Xiao Huang. 2021. "Using Mobile Device Data to Track the Effects of the COVID-19 Pandemic on Spatiotemporal Patterns of National Park Visitation" *Sustainability* 13, no. 16: 9366. https://doi.org/10.3390/su13169366

Ladyzhets, B. (2021, June 16). We investigated whether digital contact tracing actually worked in the US. MIT Technology Review. https://www.technologyreview.com/2021/06/16/1026255/us-digital-contact-tracing-exposure-notification-analysis/.

Lai, S., zu Erbach-Schoenberg, E., Pezzulo, C., Ruktanonchai, N. W., Sorichetta, A., Steele, J., Li, T., Dooley, C.A., & Tatem, A. J. (2019). *Exploring the use of mobile phone data for national migration statistics. Palgrave communications, 5*(1), 1-10.

Liu, L., Gao, X., Zhuang, J., Wu, W., Yang, B., Cheng, W., Xiao, P., Yao, X., & Deng, O. (2020). *Evaluating the lifestyle impact of China's rural housing land consolidation with locational big data: A study of Chengdu. Land use policy, 96*(104623).

Lu, X., Bengtsson, L., Holme, P. (2012). Predictability of population displacement after the 2010 Haiti earthquake. Proc. Natl. Acad. Sci. *109*(29) https://www.pnas.org/doi/full/10.1073/pnas.1203882109

Merrill, N. H., Atkinson, S. F., Mulvaney, K. K., Mazzotta, M. J., & Bousquin, J. (2020). Using data derived from cellular phone locations to estimate visitation to natural areas: An application to water recreation in New England, USA. *PloS one*, *15*(4), e0231863. https://doi.org/10.1371/journal.pone.0231863

Milusheva, S., Lewin, A., Begazo Gomez, T., Matekenya, D., & Reid, K. (2021). Challenges and opportunities in accessing mobile phone data for COVID-19 response in developing countries. *Data & Policy, 3*(E20). doi:10.1017/dap.2021.10

Monz C, Mitrovich M, D'Antonio A, Sisneros-Kidd A. (2019). Using Mobile Device Data to Estimate Visitation in Parks and Protected Areas: An example from the nature reserve of Orange County, California. *Journal of Park and Recreation Administration*. 37(4).

National Academies of Sciences, Engineering, and Medicine. (2021). Use of Vehicle Probe and Cellular GPS Data by State Departments of Transportation. Washington, DC: The National Academies Press. https://doi.org/10.17226/26094.

Office for National Statistics, (2016, July 4). ONS working paper series no 8 – Statistical uses for mobile phone data: literature review. Last accessed 5/25/2022 https://www.ons.gov.uk/methodology/methodologicalpublications/generalmethodology/onsworkingpaperseries/onsmethodologyworkingpaperseriesno8statisticalusesformobilephonedataliteraturereview

Pastor-Escuredo, D. and others (2014). "Flooding through the Lens of Mobile Phone Activity" Last accessed 12/6/2021 http://arxiv.org/ftp/arxiv/papers/1411/1411.6574.pdf

President's Council of Advisors on Science and Technology. (2014, May). Report to the President. Big Data and Privacy: A Technological Perspective. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_big_data_and_privacy_-_may_2014.pdf

Silm, S., & Ahas, R. (2014). The temporal variation of ethnic segregation in a city: Evidence from a mobile phone use dataset. *Social Science Research, 47*, 30-43.

Simmons, A. (2022, Oct 5). Cell Tower Range: How far do they reach? Dgtl Infra. https://dgtlinfra.com/cell-tower-range-how-far-reach/

Song, G., Bernasco, W., Liu, L., Xiao, L., Zhou, S., & Liao, W. (2019). Crime feeds on legal activities: Daily mobility flows help to explain thieves' target location choices. *Journal of Quantitative Criminology, 35*(4).

Sterly, H., Hennig, B., & Dongo, K. (2013). “Calling Abidjan”–Improving population estimations with mobile communication data (IPEMCODA). *Social and Economic Development, 10*, 108-114.

Streetlight. (2020. *Our methodology and data sources*. https://learn.streetlightdata.com/methodology-data-sources-white-paper

Trasberg, T., Soundararaj, B., & Cheshire, J. (2021). Using Wi-Fi probe requests from mobile phones to quantify the impact of pedestrian flows on retail turnover. *Computers, Environment and Urban Systems*, *87*(101601).

United Nations. (2019, November). *Handbook on The Use of Mobile Phone Data for Official Statistics*. Last accessed 12/6/2021 https://unstats.un.org/bigdata/task-teams/mobile-phone/MPD%20Handbook%2020191004.pdf

Valentino-DeVries, J., Singer, N., Keller, M. H., & Krolik, A.(2018, Dec. 25). Your apps know where you were last night, and they’re not keeping it secret. The New York Times. https://www.sfgate.com/business/article/Your-apps-know-where-you-were-last-night-and-13489658.php

Vanhoof M, Reis F, Ploetz T, Smoreda Z. Assessing the quality of home detection from mobile phone data for official statistics. Journal of Official Statistics. 2018 Dec 1; 34(4):935–60

Vespe, M., Iacus, S., Santamaria, C., Sermi, F., & Spyratos, S. (2021). On the use of data from multiple mobile network operators in Europe to fight COVID-19. *Data & Policy*, *3*(E8). doi:10.1017/dap.2021.9

Vlag, P., Durnell, U., & Malmros, J. (2022) Quality Improvement in Mobile Phone Position Data: A collaboration model between SCB and an operator. Presented at 2022 VILNIUS. https://q2022.stat.gov.lt/scientific-information/papers-presentations/session-38

Williams, D. & Rabinovitch, A. (2020, April 22*). Israel suspends cellphone-tracking for coronavirus quarantine enforcement*. Reuters. https://www.reuters.com/article/health-coronavirus-israel-police-idINKCN2242JF4

Williams, S., Xu, W., Tan, S. B., Foster, M. J., & Chen, C. (2019). Ghost cities of China: Identifying urban vacancy through social media data. *Cities, 94*, 275-285.

Yao, Y., Liu, X., Li, X., Zhang, J., Liang, Z., Mai, K., & Zhang, Y. (2017). Mapping fine-scale population distributions at the building level by integrating multisource geospatial big data. *International Journal of Geographical Information Science, 31*(6), 1220-1244.

Yu, H., Russell, A., Mulholland, J., Huang, Z. (2018). Using cell phone location to assess misclassification errors in air pollution exposure estimation. Environmental Pollution, 233:261–6. https://doi.org/10.1016/j.envpol.2017.10.077

Zhou, X., Lang, W., Yeh, A. G., & Niu, X. (2021). Understanding the compactness of employment activities in high-density cities through cellphone location data*. Environment and Planning B: Urban Analytics and City Science, 48*(6), 1398-1413.

Zhou, X., Yeh, A.G.O., Li, W., et al. (2018) A commuting spectrum analysis of the jobs–housing balance and self-containment of employment with mobile phone location big data. *Environment and Planning B: Urban Analytics and City Science* 45: 434–451.